AD-A040 272 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER --ETC F/G 9/2
CONVERGENCE AND COMPLEXITY OF INTERPOLATORY-NEWTON ITERATION IN--ETC(U)
APR 77 J F TRAUB, H WOZNIAKOWSKI N00014-76-C-0370

UNCLASSIFIED NL

1 OF 1
AD
A040272

END
DATE
FILMED
6-77

(12)

CONVERGENCE AND COMPLEXITY OF
INTERPOLATORY-NEWTON ITERATION
IN A BANACH SPACE

J. F. Traub
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

H. Woźniakowski
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(Visiting from the University of Warsaw)

DEPARTMENT
of
COMPUTER SCIENCE

Carnegie-Mellon University

12

CONVERGENCE AND COMPLEXITY OF INTERPOLATORY-NEWTON ITERATION IN A BANACH SPACE

J. F. Traub
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

H. Woźniakowski
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(Visiting from the University of Warsaw)

April 1977

This research was supported in part by the National Science Foundation under Grant MCS75-222-55 and the Office of Naval Research under Contract N00014-76-C-0370, NR 044-422.

ABSTRACT

The class of Interpolatory-Newton iterations is defined and analyzed for the computation of a simple zero of a non-linear operator in a Banach space of finite or infinite dimension. Convergence of the class is established.

The concepts of "informationally optimal class of algorithms" and "optimal algorithm" are formalized. For the multivariate case, the optimality of Newton iteration is established in the class of one-point iterations under an "equal cost assumption ".

1. INTRODUCTION

In Traub and Woźniakowski [76c] we investigate the class of direct interpolatory iterations I_n for a simple zero of a non-linear operator in a Banach space of finite or infinite dimension. The solution of a "polynomial operator equation" is required at each step. In this paper we consider the solution of this polynomial operator equation by a certain number of Newton iteration steps. We call this the class of Interpolatory-Newton iterations IN_n. We analyze the convergence and complexity of this class.

Traub and Woźniakowski [76c] show that the radius of the ball of convergence of I_n can grow with n. Since IN_n uses Newton iteration as its "inner process" its convergence characteristics are similar to Newton iteration (Traub and Woźniakowski [77]) and the convergence is only "local". A "type of global convergence" is established for a certain class of operators.

The complexity analysis of IN_n requires some new complexity concepts. We formalize the idea of "optimal algorithm". Under an "equal cost assumption" (and one additional reasonable assumption) we establish the optimality of I_3 for scalar problems and the optimality of $IN_2 \equiv I_2$ (Newton iteration) for multivariate problems. However, if the equal cost assumption is violated a high order iteration is optimal.

We summarize the results of this paper. Convergence of the class of iterations is established in Section 3. General complexity results are obtained in Section 4 and used to establish the optimality results of Section 5. In the final section we analyze a class of problems for which Newton iteration is not optimal.

2. INTERPOLATORY-NEWTON ITERATION IN_n

In Traub and Woźniakowski [76c] we consider interpolatory iteration I_n for the solution of the non-linear operator equation

$$(2.1) \quad F(x) = 0$$

where $F: D \subset B_1 \to B_2$ and B_1, B_2 are real or complex Banach spaces of dimension N, $N = \dim(B_1) = \dim(B_2)$, $1 \le N \le +\infty$. The interpolatory iteration I_n is defined as follows. Let x_i be an approximation to the simple solution α and let w_i be the interpolatory polynomial of degree $\le n-1$ such that

$$(2.2) \quad w_i^{(j)}(x_i) = F^{(j)}(x_i), \quad j = 0,1,\ldots,n-1$$

where $n \ge 2$. The next approximation x_{i+1}^* is a zero of w_i, $w_i(x_{i+1}^*) = 0$, with a certain criterion of its choice. Note that for $n = 2$ we get Newton iteration since $x_{i+1}^* = x_i - F(x_i)^{-1}F(x_i)$. The degree of w_i is, in general, equal to $n-1$ and for $n \ge 3$ we get a "polynomial operator equation" for x_{i+1}^*. There are a number of ways for dealing with the problem of solving this equation numerically. In this paper we will approximate x_{i+1}^* by applying a number of Newton iterations to the equation $w_i(x) = 0$. Let

$$
(2.3) \quad
\begin{aligned}
z_0 &= x_i \\
z_{j+1} &= z_j - w_i'(z_j)^{-1} w_i(z_j), \quad j = 0,1,\ldots,k-1 \\
x_{i+1} &= z_k
\end{aligned}
$$

where $k = \lceil \log_2 n \rceil$.

We shall call the iteration constructing $\{x_i\}$ by (2.3) the interpolatory-Newton iteration IN_n. This is a one-point stationary iteration without memory in the sense of Traub [64]. Note that for $n = 2$, (2.3) reduces to Newton iteration and $\mathrm{IN}_2 = \mathrm{I}_2$. To compute z_{j+1} one must solve the linear equation $w_i'(z_j)(z_j - z_{j+1}) = w_i(z_j)$. We do not specify what algorithm is used to solve this linear equation. In fact, IN_n is the name of a <u>class</u> of iterations which use the same information (2.2), and perform k Newton steps to solve $w_i(x) = 0$ but they can differ in the algorithm used to solve the linear equaton. <u>For example, by Newton iteration we mean any iteration which produces $x_{i+1} = x_i - F'(x_i)^{-1}F(x_i)$ no matter what algorithm is used to compute x_{i+1}.</u>

Some properties of the interpolatory iteration I_n, $n \geq 2$, will be used to establish the convergence of IN_n. Let α be the simple solution of $F(x) = 0$ and $J = \{x: \|x-\alpha\| \leq \Gamma\}$. Define

$$(2.4) \quad A_j = A_j(\Gamma) = \sup_{x \in J} \left\| F'(\alpha)^{-1} \frac{F^{(j)}(x)}{j!} \right\|, \quad j = 2,3,\ldots$$

whenever $F^{(j)}$ exists.

From theorem 2.1 in Traub and Woźniakowski [76c] directly follows

<u>Theorem 2.1</u>

If F is twice differentiable in J and

$$(2.5) \quad A_2 \Gamma \leq \frac{1}{4},$$

$$(2.6) \quad x_i \in J$$

then the next approximation x_{i+1}^* constructed by Newton iteration satisfies

$$(2.7) \quad \begin{aligned} \|x_{i+1}^* - \alpha\| &\leq \frac{A_2}{1 - 2A_2\|x_i - \alpha\|} \|x_i - \alpha\|^2 \leq \frac{1}{2}\|x_i - \alpha\|, \\ x_{i+1}^* - \alpha &= \frac{1}{2} F'(\alpha)^{-1} F''(\alpha)(x_i - \alpha)^2 + o(\|x_i - \alpha\|^2). \end{aligned}$$

■

Furthermore, from theorem 2.1 in Traub and Woźniakowski [76c] follows

Theorem 2.2

If F is n-times differentiable, $n \geq 3$, in J, and

$$(2.8)\quad \frac{nA_n\Gamma^{n-1}}{1-A_2\Gamma} < \left(\frac{2}{3}\right)^{n-1} ,$$

$$(2.9)\quad x_i \in J$$

then the polynomial w_i has a unique zero in $J_1 = \{x: \|x-\alpha\| \leq \frac{1}{2}\Gamma\}$ and defining x_{i+1}^* as the zero of w_i in J_1 we get

$$(2.10)\quad \|x_{i+1}^*-\alpha\| \leq \frac{A_n(1+\|x_{i+1}^*-\alpha\|/\|x-\alpha\|)^n}{1-A_2\|x_{i+1}^*-\alpha\|} \cdot \|x_i-\alpha\|^n \leq \frac{1}{2}\|x_i-\alpha\| ,$$

$$(2.11)\quad x_{i+1}^*-\alpha = \frac{(-1)^n}{n!} F'(\alpha)^{-1}F^{(n)}(\alpha)(x_i-\alpha)^n + o(\|x_i-\alpha\|^n). \quad \blacksquare$$

3. CONVERGENCE OF INTERPOLATORY-NEWTON ITERATION

We study the convergence of IN_n for $n \geq 3$. Let $e_i = \|x_i - \alpha\|$, $\forall_i$.

<u>Theorem 3.1</u>

If F is n-times differentiable, $n \geq 3$, in J and

(3.1) $\quad 0 \leq \tilde{A}_2\Gamma \leq 1/5$

where
$$\tilde{A}_2 = \frac{A_2 + \frac{n(n-1)}{2}A_n(2\Gamma)^{n-2}}{1 - A_2\Gamma - nA_n\left(\frac{3}{2}\right)^{n-1}\Gamma^{n-1}},$$

(3.2) $\quad x_0 \in J$

then the sequence $\{x_i\}$ constructed by the interpolatory-Newton iteration IN_n is well-defined and

(3.3) $\quad x_i \in J, \forall i,$

(3.4) $\quad \lim_i x_i = \alpha, \quad e_{i+1} \leq \{\frac{1}{2} + \frac{3}{2}\left(\frac{1}{2}\right)^k\}e_i, \quad \forall i,$

(3.5) $\quad e_{i+1} \leq C_{i,n}\, e_i^n$ where

$$C_{i,n} = \left(1+\frac{e^*_{i+1}}{e_i}\right)^n\left(\frac{A_n}{1-A_2e^*_{i+1}} + [\tilde{A}_2(1+H_i)]^{2^k-1}\left[\left(1+\frac{e^*_{i+1}}{e_i}\right)e_i\right]^{2^k-n}\right)$$

for $\quad e^*_{i+1} = \|x^*_{i+1} - \alpha\|$, $H_i = O(e_i)$ and $0 \leq H_i \leq \frac{5}{2}$, $k = \lceil \log_2 n \rceil$,

$\lim_i C_{i,n} = A_n + \delta\tilde{A}_2^{\,n-1}$ where $\delta = 0$ if $2^k = n$, and $\delta = 1$ if $2^k = n$,

$$(3.6)\quad x_{i+1}-\alpha = F_n(x_i-\alpha)^n + b_{i,k} + o(\|x_i-\alpha\|^n)$$

where

$$b_{i,1} = F_2(x_i-\alpha)^2$$

$$b_{i,j+1} = F_2\, b_{i,j}^2\ , \quad j = 1,2,\ldots,k-1$$

and

$$F_j = \frac{(-1)^j}{j!} F'(\alpha)^{-1} F^{(j)}(\alpha) \quad \text{for } j = 2 \text{ and } n. \quad \blacksquare$$

<u>Proof</u>

Assume by induction that $x_i \in J$. We want to show that the interpolatory-Newton iteration is well-defined, i.e., $w_i'(z_j)$ is invertible. First we shall prove that $w'(x)$ is invertible for $x \in J$ and next that $z_j \in J$. Denote

$$(3.7)\quad F^{(j)}(x) - w_i^{(j)}(x) = R_n^{(j)}(x;x_i) \quad \text{for } x \in J,\ j = 0,1,2,$$

where

$$\|F'(\alpha)^{-1} R_n^{(j)}(x;x_i)\| \le j!\binom{n}{j} A_n \|x-x_i\|^{n-j},$$

see Rall [69, p.124]. Since

$$w_i'(x) = F'(x) - R_n'(x;x_i) = F'(\alpha)[I + F'(\alpha)^{-1}\{F'(x) - F'(\alpha)\} - F'(\alpha)^{-1}R_n'(x;x_i)]$$

then from (3.7) and (3.1) we get for $x \in J$

$$(3.8)\quad \|F'(\alpha)^{-1}w_i'(x) - I\| \le 2A_2\|x-\alpha\| + nA_n\|x-x_i\|^{n-1} \le$$

$$\le 2A_2\Gamma + nA_n(2\Gamma)^{n-1} \le \frac{2}{5} < 1.$$

From theorem 10.1 in Rall [69, p.36] it follows that $w_i'(x)$ is invertible for any $x \in J$ and

$$(3.9)\quad \|w_i'(x)^{-1}F'(\alpha)\| \le \frac{1}{1 - 2A_2\|x-\alpha\| - nA_n\|x-x_i\|^{n-1}} .$$

Since the denominator in (3.1) is positive then

$$\frac{nA_n\Gamma^{n-1}}{1-A_2\Gamma} < \left(\frac{2}{3}\right)^{n-1}$$

and from theorem 2.2 follows that the polynomial w_i has a unique zero in $J_1 = \{x\colon \|x-\alpha\| \le \frac{1}{2}\Gamma\}$, $w(x_{i+1}^*) = 0$, and (2.10) holds. From (3.9) and (3.7) we get for $x \in J$

$$(3.10)\quad \|w_i'(x_{i+1}^*)^{-1}\frac{w_i''(x)}{2}\| \le \|w_i'(x_{i+1}^*)^{-1}F'(\alpha)\|\,\|F'(\alpha)^{-1}\frac{w_i''(x)}{2}\| \le$$

$$\le \frac{A_2 + \frac{n(n-1)}{2}A_n\|x-x_i\|^{n-2}}{1 - 2A_2\|x_{i+1}^*-\alpha\| - nA_n\|x_{i+1}^*-x_i\|^{n-1}} \le \frac{A_2 + \frac{n(n-1)}{2}A_n(2\Gamma)^{n-2}}{1 - A_2\Gamma - nA_n(\frac{3}{2}\Gamma)^{n-1}} = \tilde{A}_2 .$$

We investigate the properties of $\{z_j\}$ defined in (2.3). Recall we solve $w_i(x) = 0$ by Newton iteration. Note that $z_1 = x_i - F'(x_i)^{-1}F(x_i)$ is the Newton step applied to the equation $F(x) = 0$. From (3.1) we know that $A_2\Gamma \le 1/4$ and from theorem 2.1 we get

$$\|z_1-\alpha\| \le \frac{1}{2}\|x_i-\alpha\| .$$

Since $x_{i+1}^* \in J_1$, $\|z_1-x_{i+1}^*\| \le \Gamma$.

We prove that z_{j+1} lies in $D_j = \{x\colon \|x-x_{i+1}^*\| \le \frac{1}{2}\|z_j-x_{i+1}^*\|\} \cap J$. Let

$$(3.11)\quad w_i(x) = w_i(z_j) + w_i'(z_j)(x-z_j) + \bar{R}_2(x;z_j)$$

where $\bar{R}_2(x;y) = \int_0^1 w_i''(y+t(x-y))(x-y)^2(1-t)dt,$

for $x,y \in J$, compare with (3.7). Note that z_{j+1} is the zero of the equation

$$(3.12)\quad x = H(x) \overset{\text{df}}{=} x_{i+1}^{*} + w'(x_{i+1})^{-1}\{\bar{R}_2(x;z_j) - \bar{R}_2(x;x_{i+1}^{*})\}.$$

We show that H is contractive on D_j. From (3.10) we have for $x \in D_j$

$$\|H(x)-x_{i+1}^{*}\| \le \tilde{A}_2(\|x-z_j\|^2 + \|x-x_{i+1}^{*}\|^2) \le \tfrac{5}{2}\tilde{A}_2\|z_j-x_{i+1}^{*}\|^2 \le \tfrac{1}{2}\|z_j-x_{i+1}^{*}\|$$

due to (3.1). Furthermore

$$\|H(x)-\alpha\| \le \|x_{i+1}^{*}-\alpha\| + \|H(x)-x_{i+1}^{*}\| \le (\tfrac{1}{2}+\tfrac{1}{2})\Gamma = \Gamma.$$

Thus $H(D_j) \subset D_j$. Since $\|H'(x)\| \le 2\tilde{A}_2\|z_j-x_{i+1}^{*}\| \le 2\tilde{A}_2\Gamma < 1$ then H is contractive on D_j and z_{j+1} is the unique zero in D_j.

This proves that $x_{i+1} = z_k \in J$ and

$$\|x_{i+1}-\alpha\| \le \|x_{i+1}-x_{i+1}^{*}\| + \|x_{i+1}-\alpha\| \le \left(\tfrac{1}{2}\right)^k\|z_0-x_{i+1}^{*}\| + \tfrac{1}{2}\|x_i-\alpha\| \le$$

$$\le \left(\tfrac{3}{2}\left(\tfrac{1}{2}\right)^k + \tfrac{1}{2}\right)\|x_i-\alpha\| \le \tfrac{7}{8}\ \|x_i-\alpha\|$$

which yields (3.3) and (3.4).

Let $\tilde{e}_j = \|z_j-x_{i+1}^{*}\|$. Set $x = z_{j+1}$ in (3.12). Then

$$(3.13)\quad \tilde{e}_{j+1} \le \frac{\tilde{A}_2(1+\tilde{e}_{j+1}/\tilde{e}_j)^2}{1-\tilde{A}_2\tilde{e}_{j+1}}\tilde{e}_j^2 \le \tilde{A}_2(1+H_i)\tilde{e}_j$$

where $H_i = O(\tilde{e}_j)$ and $0 \le H_i \le 5/2$, compare (2.7). Since $\tilde{e}_j = O(e_i)$ we can write $H_i = O(e_i)$. Next from (3.13) and (2.10) we get

$$e_{i+1} = \|x_{i+1}-\alpha\| \le \|x_{i+1}-x^*_{i+1}\| + \|x^*_{i+1}-\alpha\| = \tilde{e}_k + \|x^*_{i+1}-\alpha\| \le$$

$$\le [\tilde{A}_2(1+H_i)]^{2^k-1}\|x_i-x^*_{i+1}\|^{2^k} + \frac{A_n}{1-A_2e^*_{i+1}}\left(1+\frac{e^*_{i+1}}{e_i}\right)^n e_i^n \le$$

$$\le \left(1+\frac{e^*_{i+1}}{e_i}\right)^n\left(\frac{A_n}{1-A_2e^*_{i+1}} + [\tilde{A}_2(1+H_i)]^{2^k-1}\left[\left(1+\frac{e^*_{i+1}}{e_i}\right)e_i\right]^{2^k-n}\right)e_i^n =$$

$$= C_{i,n}e_i^n .$$

Since e^*_{i+1}/e_i and H_i tend to zero then

$$\lim_i C_{i,n} = A_n + \delta\tilde{A}_2^{n-1}$$

where $\delta = 0$ if $2^k > n$ and $\delta = 0$ otherwise. Hence (3.5) holds.

Finally observe that

$$z_{j+1} - x^*_{i+1} = w_i'(x^*_{i+1})^{-1}\frac{w_i''(x^*_{i+1})}{2}(z_j-x^*_{i+1})^2 + O(\tilde{e}_j^3) =$$

$$= F'(\alpha)^{-1}\frac{F''(\alpha)}{2}(z_j-x^*_{i+1})^2 + O(e^*_{i+1}\tilde{e}_j^2 + \tilde{e}_j^3) =$$

$$= F_2(z_j-x^*_{i+1})^2 + o(\tilde{e}_j^2).$$

Thus

$$z_k - x^*_{i+1} = F_2(F_2\cdot\ldots\cdot(F_2(x_i-x^*_{i+1})^2)^2 \ldots)^2 + o(e_i^{2^k}) =$$

$$= F_2(F_2\cdot\ldots\cdot(F_2(x_i-\alpha)^2)^2 \ldots)^2 + o(e_i^{2^k}).$$

From the definition of $b_{i,k}$ in (3.6) we get

$$z_k - x^*_{i+1} = b_{i,k} + o(e_i^{2^k}).$$

From this and (2.11) we have

$$x_{i+1} - \alpha = z_k - x^*_{i+1} + x^*_{i+1} - \alpha = b_{i,k} + F_n(x_i-\alpha)^n + o(e_i^n)$$

which proves (3.6) and completes the proof of theorem 3.1. ■

Remark 3.1

It is possible to get a slightly better estimate than (3.1) although the proof is much more complicated. Note that if n is not a power of 2 then $2^k > n$ and the leading term in (3.6) is $F_n(x_i-\alpha)$ since $\|b_{i,k}\| = O(e_i^{2^k}) = o(e_i^n)$. ■

Remark 3.2

The idea of using Newton iteration to estimate a zero of an approximating non-linear operator which fits the information of F can be applied for any iterations with or without memory; see Brent [76] where Newton iteration is also used as an inner iteration. ■

Remark 3.3

Since Newton iteration is numerically stable it is relatively easy to verify the numerical stability of the interpolatory-Newton iteration IN_n under appropriate assumptions on the computed information of F, see Woźniakowski [76b]. ■

In general the interpolatory-Newton iteration converges only locally. We give conditions under which IN_n enjoys a "type of global convergence". Compare the same property for Newton iteration in Traub and Woźniakowski [77].)

Let $F(x) = \sum_{i=1}^{\infty} \frac{1}{i!} F^{(i)}(\alpha)(x_i-\alpha)^i$ be analytic in $D = \{x: \|x-\alpha\| < R\}$ and

$$(3.14)\quad \frac{\|F'(\alpha)^{-1}F^{(i)}(\alpha)\|}{i!} \le K^{i-1}$$

for $i = 2,3,\ldots$, $R \geq 1/K$.

One way to find K is to use Cauchy's formula

$$\frac{\|F'(\alpha)^{-1}F^{(i)}(\alpha)\|}{i!} \leq \frac{M}{R^i}$$

where $M = \sup_{x \in D} \|F'(\alpha)^{-1}F(x)\|$. Setting $K = \max(\frac{1}{R}, \frac{M}{R^2})$ we get $M/R \leq KR \leq (KR)^{i-1}$ which yields $M/R^i \leq K^{i-1}$.

Theorem 3.2

If F satisfies (3.14) then the interpolatory Newton converges for $x_0 \in J = \{x: \|x-\alpha\| \leq \Gamma_n\}$ where

$$\Gamma_n = \frac{x_n}{K}$$

and x_n, $0 < x_n < x_\infty$, satisfies the equation

$$(3.15) \quad 5\left(\frac{x}{(1-x)^3} + \frac{n(n-1)}{4(1-x)^2}\left(\frac{2x}{1-x}\right)^{n-1}\right) = 1 - \frac{x}{(1-x)^3} - \frac{n}{(1-x)^2}\left(\frac{3x}{2(1-x)}\right)^{n-1},$$

and $x_n \nearrow x_\infty$ where x_∞ is the smallest positive solution of $x/(1-x)^3 = 1/6$ and $x_\infty \doteq .12$. ∎

Proof

From (3.14) we get

$$\|F'(\alpha)^{-1}F^{(i)}(x)\| \leq f^{(i)}(\|x-\alpha\|)$$

where $f(x) = x/(1-Kx)$. Since $f^{(i)}(x) = i!K^{i-1}/(1-Kx)^{i+1}$ for $i \geq 2$, we have

$$A_i(\Gamma) \leq \frac{K^{i-1}}{(1-K\Gamma)^{i+1}}, \quad i = 2,3,\ldots .$$

Then (3.1) becomes

$$\tilde{A}_2\Gamma \le \left(\frac{K\Gamma}{(1-K\Gamma)^3} + \frac{n(n-1)}{4(1-K\Gamma)^2}\left(\frac{2K\Gamma}{1-K\Gamma}\right)^{n-1}\right) \Big/ \left(1 - \frac{K\Gamma}{(1-K\Gamma)^3} - \frac{n}{(1-K\Gamma)^2}\left(\frac{3K\Gamma}{2(1-K\Gamma)}\right)^{n-1}\right) = \frac{1}{5}$$

Setting $K\Gamma = x$ we get that x satisfies the equation (3.15). It is straightforward to verify that $x = x(n)$ is an increasing function of n and $\lim_n x(n) = x_\infty$ where x_∞ satisfies the equation

$$\frac{x}{(1-x)^3} = \frac{1}{6} .$$

Hence $x_\infty \doteq .12$ which proves theorem 3.2. ■

This result is especially interesting if the domain radius R is related to $\frac{1}{K}$, say $R = \frac{c_1}{K}$. Then $\Gamma_n = \frac{x_n}{K} = \frac{x_n}{c_1} R \cong \frac{0.12}{c_1} R$ and the interpolatory-Newton iteration enjoys a "type of global convergence".

4. COMPLEXITY OF ONE-POINT ITERATIONS

In this section we deal with complexity of one-point iterations. We extend some of the results of Traub and Woźniakowski [76a].

Assume that a one-point iteration φ constructs the sequence $x_{i+1} = \varphi(x_i;F)$ converging to α and satisfying

(4.1) $e_i = G_i e_{i-1}^p$, $e_i = ||x_i-\alpha||$, $i = 1,2,\ldots,K$

where p, $p > 1$, is called the order of iteration φ,

(4.2) $0 < \underline{G} \le G_i \le \bar{G} < +\infty$, $i = 1,2,\ldots,K$

and the iteration is terminated after K steps.

From (4.1) we get

(4.3) $$e_i = \left(\frac{1}{w_i}\right)^{p^i-1} e_0 \text{ where } \frac{1}{w_i} = \left(G_1^{p^{i-1}} G_2^{p^{i-2}} \cdot\ldots\cdot G_i\right)^{\frac{1}{p^{i-1}}} e_0.$$

Note that $(e_0 w_i)^{1-p}$ is the geometric mean of the $G_1,G_2,\ldots,G_i$. Furthermore $e_i < e_0$ iff $w_i > 1$. From (4.2) we get

(4.4) $$\frac{1}{\underline{w}} = \underline{G}^{\frac{1}{p-1}} e_0 \le \frac{1}{w_i} \le \bar{G}^{\frac{1}{p-1}} e_0 = \frac{1}{\bar{w}}.$$

We shall assume that $\bar{w} > 1$.

For a given ϵ', $0 < \epsilon' < 1$, let K be the smallest index for which $e_K \le \epsilon' e_0$. Define $\epsilon < \epsilon'$ so that

(4.5) $e_K = \epsilon e_0$.

Let comp = comp(φ,F) be the total cost of finding x_K. Assume that the cost of the ith iterative step does not depend on the index i; we denote it by c = c(φ;F). Then

(4.6) comp = cK.

From (4.3) and (4.5) we get

$$\left(\frac{1}{w_K}\right)^{p^K-1} = \varepsilon \text{ and } K = \frac{g(w_K)}{\log p}$$

where

(4.7) $g(w) = \lg(1 + \frac{t}{\lg w}), \quad t = \lg 1/\varepsilon.$

We take all logarithms for the remainder of this paper to base 2. Then from (4.6) and (4.7) we get

(4.8) $\text{comp} = z\, g(w_K)$

where

(4.9) $z = \frac{c}{\lg p}$

is called the <u>complexity index</u>.

Since g(w) is a monotonically decreasing function, (4.4) gives bounds on complexity

(4.10) $zg(\underline{w}) \le \text{comp} \le zg(\bar{w})$.

Note that as $\varepsilon \to 0$, $g(w) \cong \lg t$ and comp $\cong z\lg t$. If we assume that

(4.11) $2 \le \bar{w} \le \underline{w} \le t$

then (4.10) becomes

$$(4.12) \quad z(\lg t - \lg\lg t) \le \text{comp} \le z \lg(1+t);$$

see theorem 3.1 in Traub and Woźniakowski [76a]. In this case the complexity index is a good measure of complexity.

We want to minimize the total cost of finding x_K. More precisely, for a given operator F we want to find an iteration Φ with minimal complexity. Since we do not know the value $g(w_K)$ in (4.8) we are not able to minimize complexity. However, if (4.11) holds or ε is small enough then minimal complexity is approximated for an iteration with minimal complexity index. So we wish to find an iteration Φ which for a given problem has as small a complexity index as possible.

The complexity index is given by $z = c/\lg p$ where c is the cost per one iterative step and p is the order of an iteration. Assume that an iteration φ_n uses the <u>standard information</u> (see Woźniakowski [75a]) $\mathfrak{N}_n = \mathfrak{N}_n(x;F) = \{F(x), F'(x), \ldots, F^{(n-1)}(x)\}$. The cost $c = c(\varphi_n, \mathfrak{N}_n)$ consists of the information complexity $u = u(F, \mathfrak{N}_n)$ which is the cost of computing $\mathfrak{N}_n(x;F)$ and the combinatory complexity $d = d(\varphi_n)$ which is the cost of combining information and producing the next approximation. Then $c = u+d$ and (4.9) can be rewritten as

$$(4.13) \quad z(\varphi_n) = \frac{u(F,\mathfrak{N}_n) + d(\varphi_n)}{\lg p(\varphi_n)} .$$

We want to find an iteration φ_n which minimizes (4.13). Let

$$(4.14) \quad z_n(F) = \inf_{\varphi_n \in \Phi_n} z(\varphi_n),$$

$$(4.15) \quad z(F) = \inf_{n \ge 2} z_n(F),$$

where Φ_n is the class of one-point iterations using the standard information with n evaluations. (Of course $IN_n \in \Phi_n, \forall n$.) We need lower and upper bounds on $z_n(F)$ and $z(F)$. Note that upper bounds can be obtained by the complexity index of any iteration φ_n, e.g., by the complexity index of the interpolatory-Newton iteration IN_n. We shall deal with this in Section 5.

Throughout the rest of this paper we assume that the dimension of the problem is finite, $\underline{N < +\infty}$.

To find lower bounds on $z_n(F)$ and $z(F)$ we need a lower bound on $z(\varphi_n)$. Note that

$$(4.16)\quad u(F;\mathfrak{N}_n) = \sum_{i=0}^{n} c(F^{(i)})$$

where $c(F^{(i)})$ denotes the "cost" of computing $F^{(i)}(x)$. In the "cost" one can include all the costs of computing $F^{(i)}$ including the cost of all arithmetic operations needed to compute $F^{(i)}$, the cost of variable data access, the cost of subroutine calls, etc. For the sake of simplicity we assume that the cost of one arithmetic operation is taken as unity.

In general $F^{(j)}(x)$ requires $N\binom{N+j-1}{j}$ different data for its representation. The total number of data in $\mathfrak{N}_n$ is equal to

$$(4.17)\quad d_{N,n} = N\sum_{j=0}^{n-1}\binom{N+j-1}{j} = N\binom{N+n-1}{n-1}.$$

For almost all problems the information cost $u(F;\mathfrak{N}_n)$ depends linearly on $d_{N,n}$. To make this more precise we introduce

<u>Definition 4.1</u> (Functional Independence Assumption)

We say F satisfies the <u>functional independence assumption</u> if there exists a positive constant c_F such that

$$(4.18)\quad u(F;\mathfrak{N}_n) \ge c_F\, d_{N,n}, \quad \forall n.$$ ∎

If F depends on all different data in $\mathfrak{N}_n$ then, of course, c_F is at least equal to unity. However it can happen due to a special property of F (like symmetry of some $F^{(i)}$) that the information cost $u(F;\mathfrak{N}_n)$ is less than $d_{N,n}$. In the functional independence assumption we need not specify the value of c_F as long as c_F is positive.

We estimate the combinatorial complexity $d(\varphi_n)$. Any iteration φ has to use every piece of data at least once as well as the current approximation to α. Thus

$$(4.19)\quad d(\varphi_n) \ge d_{N,n}.$$

Since $p(\varphi_n) \le n$ (see Traub and Woźniakowski [76b]) for any iteration φ we find

$$(4.20)\quad z(\varphi_n) = \frac{u(F;\mathfrak{N}_n) + d(\varphi_n)}{\lg p(\varphi_n)} \ge (c_F+1)\,\frac{d_{N,n}}{\lg n} \ge (c_F+1)\,V_1$$

where $V_1 = \frac{3}{\log 3}$ for $N = 1$ and $V_1 = N(N+1)$ for $N \ge 2$. Thus, we proved

<u>Theorem 4.1</u>

For the class of one-point iterations which use the standard information of F with n evaluations where F satisfies the functional independence assumption, the minimal complexity indexes $z_n(F)$ and $z(F)$ are bounded below by

$$(4.21)\quad z_n(F) \ge (c_F+1)\,\frac{d_{N,n}}{\lg n},$$

$$(4.22)\quad z(F) \ge (c_F+1)\begin{cases} \frac{3}{\lg 3} & \text{for } N = 1 \\ N(N+1) & \text{for } N \ge 2. \end{cases}$$ ∎

5. OPTIMALITY THEOREMS

In this section we deal with the complexity of the interpolatory-Newton iteration IN_n, $n \geq 2$. Recall that under the assumptions of theorem 2.1 for $n = 2$ and theorem 3.1 for $n \geq 3$ the iteration IN_n constructs the sequence $\{x_i\}$ such that

$$(5.1) \quad e_i = G_i e_{i-1}^n$$

where $e_i = \|x_i - \alpha\|$ and $G_i \leq \bar{G}$ for

$$(5.2) \quad \bar{G} = \bar{G}(n) = \begin{cases} A_2/(1 - 2A_2\Gamma) & \text{for } n = 2 \\ (1+q)^n\left(\dfrac{A_n}{1-A_2\Gamma/2} + \left(\dfrac{7}{2}\tilde{A}_2\right)^{2^k-1}[(1+q)\Gamma]^{2^k-n}\right) & \text{for } n > 2 \end{cases}$$

where $\tilde{A}_2$ is defined by (3.1), $q = \frac{1}{2} + \frac{3}{2}(\frac{1}{2})^k$ and $k = \lceil \lg n \rceil$; see (2.7) and (3.5).

Furthermore, recall (2.7) and (3.6),

$$(5.3) \quad x_{i+1} - \alpha = F_n(x_i - \alpha)^n + b_{i,k} + o(\|x_i - \alpha\|^n)$$

where $F_n = (-1)^n \frac{1}{n!} F'(\alpha)^{-1} F^{(n)}(\alpha)$ and $b_{i,k}$ given by (3.6) is omitted for $n = 2$. From (5.3) it is reasonable to assume that

$$(5.4) \quad e_{i+1} \geq \underline{G}\, e_i^n \qquad \text{for } i = 0,1,\ldots,K,$$

where $\underline{G} = \underline{G}(n)$ is a positive number.

<u>Theorem 5.1</u>

If $\bar{G}\,(2\Gamma)^{n-1} \leq 1$ then

$$(5.5) \quad \mathrm{comp}(IN_n) \leq z(IN_n)\,\lg(1+t), \quad t = \lg \frac{1}{\varepsilon}.$$

If $\underline{G}\,(t\Gamma)^{n-1} \geq 1$ then

(5.6) $\mathrm{comp}(IN_n) \geq z(IN_n)(\mathrm{lgt}-\mathrm{lglgt})$. ■

Proof

From (4.4), (4.11) and (4.12) it is enough to assure that $\frac{1}{\bar{w}} = \bar{G}^{1/(n-1)}\Gamma \leq \frac{1}{2}$ and $\frac{1}{\underline{w}} = \underline{G}^{1/(n-1)}\Gamma \geq t^{-1}$. This is equivalent to $\bar{G}(2\Gamma)^{n-1} \leq 1$ and $\underline{G}(t\Gamma)^{n-1} \geq 1$ which hold due to the assumptions. ■

Remark 5.1

To assure convergence of the interpolatory-Newton iteration IN_n, $n \geq 3$, we assumed in theorem 3.1 $\tilde{A}_2\Gamma \leq \frac{1}{5}$. To get a good upper bound on complexity we have to strengthen this inequality to $\bar{G}(2\Gamma)^{n-1} \leq 1$. It may be shown that $\bar{G}(2\Gamma)^{n-1} \leq 1$ for $n = 2^k$ implies $\tilde{A}_2\Gamma \leq \frac{2}{21}$. (However both inequalities $\tilde{A}_2\Gamma \leq \frac{1}{5}$ and $\tilde{A}_2\Gamma \leq \frac{2}{21}$ seem to be slightly overestimated.)

Note that in general a stronger condition is needed to get "good complexity" than merely to assure convergence. In Traub and Woźniakowski [77] we showed that for Newton iteration it is necessary to assume $A_2\Gamma < \frac{1}{3}$ for convergence and $A_2\Gamma \leq \frac{1}{4}$ for "good complexity". Note that for $n = 2$, $\bar{G}2\Gamma \leq 1$ is equivalent to $A_2\Gamma \leq \frac{1}{4}$. ■

We discuss the complexity index $z(IN_n)$ of the interpolatory-Newton iteration. Recall that the next approximation is obtained by $k = \lceil \lg n \rceil$ Newton steps applied to $w_i(x) = 0$. It may be shown that the total number of arithmetic operations sufficient to perform one iterative step of IN_n is equal to

$$(5.7)\quad d(IN_n) = \begin{cases} O\left(N^{\beta}\lceil \lg n\rceil + N^2\binom{N+n-1}{n-1}(\lceil \lg n\rceil - 1)\right) & \text{for } N \geq 2 \\ 5\cdot n\lceil \lg n\rceil + O(1) & \text{for } N = 1 \end{cases}$$

where the total number of arithmetic operations necessary to solve a system of N linear equations is $O(N^{\beta})$, $\beta \le 3$.

Example 5.1

Consider the iteration IN_3 for the scalar case. We solve $w_i(x) = 0$ by applying $k = \lceil lg\ 3 \rceil = 2$ Newton steps. Then

$$x_{i+1} = z_2 = x_i - \frac{F(x_i)}{F'(x_i)} - \frac{F''(x_i)}{2F'(x_i)} \left[\frac{F(x_i)}{F'(x_i)}\right]^2 \frac{1}{1-\rho_i}$$

where $\rho_i = \frac{F''(x_i)}{F'(x_i)} \frac{F(x_i)}{F'(x_i)}$. Thus the combinatory cost $d(IN_3) = 9$.

Note that $\rho_i = O(e_i)$. It may be shown that $\rho_i = 0$ does not affect the order of iteration. Therefore one can define the next approximation $\tilde{x}_{i+1}$ as

$$(5.8) \quad \tilde{x}_{i+1} = x_i - \frac{F(x_i)}{F'(x_i)} - \frac{F''(x_i)}{2F'(x_i)} \left[\frac{F(x_i)}{F'(x_i)}\right]^2 .$$

It is easy to show that $\tilde{x}_{i+1} = \tilde{w}_i(0)$ where $\tilde{w}_i^{(j)}(F(x_i)) = g^{(j)}(F(x_i))$ for $j = 0,1,2$ where $\tilde{w}_i$ is a polynomial of degree at most 2 and $g(x) = F^{-1}(x)$ is the inverse function of F, see Kung and Traub [74]. The iteration which constructs $\tilde{x}_{i+1}$ is called the <u>inverse interpolatory iteration</u> $\tilde{I}_3$ and $d(\tilde{I}_3) = 7$. ∎

A similar upper bound, $O(n\ lg\ n)$, on the combinatory complexity for the scalar case for inverse interpolatory iteration has been obtained by Brent and Kung [76].

We believe there exist no iterations with essentially less combinatory cost than that given by (5.7). We propose

Conjecture 5.1

The combinatory complexity $d(\varphi_n)$ of any iteration with maximal order, $p(\varphi_n) = n$, has to be at least

$$c_1 N^2 \binom{N+n-1}{n-1} \log n \quad \text{for } n \geq 3 \quad \text{and} \quad c_2 N^\beta \quad \text{for } n = 2$$

for positive c_1 and c_2 independent of n and N. ■

We turn to the problem of bounds on the minimal complexity index $z_n(F)$ and $z(F)$, see (4.14) and (4.15).

Definition 5.1

Let

$$z(F) = z_{n^*}(F)$$

for some integer n^*. Then we say n^* is the optimal information number with respect to z(F) (or the optimal information number) and $\mathfrak{N}_{n^*}$ is the optimal information set (among one-point standard information). An iteration ϕ_{n^*} (or a class of iterations ϕ_{n^*}) is said to be informationally optimal. ■

From (4.14) and (4.15) we get

$$z_n(F) \leq \frac{\sum_{i=0}^{n-1} c(F^{(i)}) + d(IN_n)}{\lg n},$$

$$(5.9) \qquad z(F) \leq \min_{n\geq 2} \left(\frac{\sum_{i=0}^{n-1} c(F^{(i)}) + d(IN_n)}{\lg n} \right).$$

To get further estimates on the complexity index and to find the optimal value of n we must specify a relation among the $c(F^{(i)})$. Let $c(F)$ denote the cost of evaluating $F(x)$ and assume that each new piece of data in $F(x), F'(x), \ldots$ costs the same number of arithmetic operations. For most problems $c(F)$ is proportional to N.

Definition 5.2 (Equal cost assumption)

We say F satisfies the equal cost assumption if

$$(5.10)\quad c(F^{(i)}) = \binom{N+i-1}{i} c(F) \quad \text{for } i = 1,2,\ldots \qquad \blacksquare$$

Note that the equal cost assumption implies the functional independence assumption but that the converse does not hold.

From (5.10) we get

$$(5.11)\quad \frac{\sum_{i=0}^{n-1} c(F^{(i)})}{\lg n} = \frac{\binom{N+n-1}{n-1}}{\lg n} c(F) \geq \begin{cases} (N+1)\ c(F) & \text{for } N \geq 2 \\ \frac{3}{\lg 3}\ c(F) & \text{for } N = 1. \end{cases}$$

Theorem 5.2

If F satisfies the equal cost assumption then the minimal complexity index $z(F)$ satisfies

(i) for $N = 1$

$$(5.12)\quad \frac{3}{\lg 3} c(F) + \frac{3}{\lg 3} \leq z(F) \leq \frac{3}{\lg 3} c(F) + \frac{7}{\lg 3},$$

$$(5.13)\quad z(F) = z_3(F) \quad \text{whenever} \quad c(F) \geq H_1 = 23$$

which means that the optimal information number $n^* = 3$,

(ii) for $N \geq 2$

$$(5.14)\quad (N+1)[c(F)+N] \leq z(F) \leq (N+1)[c(F) + d(IN_2)/(N+1)],$$

$$(5.15)\quad z(F) = z_2(F) \quad \text{whenever} \quad c(F) \geq H_N$$

where $H_n = [d(IN_2)/(N+1) - N(N+2)/(2 \lg 3)]/[(N+2)/(2 \lg 3) - 1]$.

which means that the optimal information number $n^* = 2$. ■

Proof

Let $N = 1$. Note that (5.12) directly follows from (4.22), (5.11) and the fact that $d(\tilde{I}_3) = 7$. We show that $z_3(F) \leq z_n(F)$ for any n. In fact

$$z_3(F) \leq \frac{3}{\lg 3} c(F) + \frac{7}{\lg 3} \leq \min_{n \neq 3} \frac{n}{\lg n}(c(F)+1) = 2(c(F)+1) \leq \min_{n \neq 3} z_n(F).$$

This yields $c(F) \geq \left(\frac{7}{\lg 3} - 2\right)\Big/\left(2 - \frac{3}{\lg 3}\right) \doteq 22.5$ which holds due to (5.13). Hence $z(F)$ is minimized for $n = 3$.

Let $N \geq 2$. The lefthand side of (5.14) follows from (4.22) since $Nc_F = c(F)$. The righthand side of (5.14) is the complexity index of Newton iteration. To prove (5.15) observe

$$z_2(F) \leq (N+1)[c(F) + d(IN_2)/(N+1)] \leq \min_{n \geq 3} \frac{\binom{N+n-1}{n-1}}{\lg n}(c(F)+N) =$$

$$= \binom{N+2}{2}(c(F)+N)/\lg 3 \leq \min_{n \geq 3} z_n(F),$$

which is equivalent to $c(F) \geq H_N$. This holds due to (5.15). Hence $z(F)$ is minimized for $n = 2$. ■

Remark 5.2

If Gauss elimination is used for solving linear equations then the combinatory cost $d(IN_2)$ of Newton iteration is equal to $d(IN_2) = \frac{2}{3} N^3 + O(N^2)$. (Note that we count all arithmetic operations.) Then H_n in (5.15) satisfies $H_N \cong \left(4 \frac{\lg 3}{3} - 1\right) N \cong 1.1N$. If the Strassen algorithm is used to solve the linear equations then $d(IN_2) = O(N^{\beta})$, $\beta = \lg 7 \doteq 2.81$ and $H_N < 0$ for large N. This means the assumption $c(F) \geq H_n$ is not restrictive in this case. ■

Theorem 5.2 states that the optimal information number is achieved for small n, $n^* = 3$ in the scalar case and $n^* = 2$ in the multivariate case, when the equal cost assumption holds and c(F) is reasonably large. However, we shall see in Section 6 that the optimal information number need not be small if the equal cost assumption does not hold.

We know the optimal value of n when F satisfies the equal cost assumption. We seek an iteration whose complexity index $z(\varphi_n)$ is equal to z(F). If c(F) is large both lower and upper bounds for z(F) in (5.12) and (5.14) are tight. Since the righthand side of (5.12) is close to the complexity index of IN_3 and the righthand side of (5.14) is the complexity index of Newton iteration IN_2, we see that any iterations in the class IN_3 and IN_2 are close to optimal among all one-point iterations in the scalar and multivariate cases, respectively. Compare with theorem 4.2 in Kung and Traub [74] where the scalar case is discussed.

We formalize the idea of optimal algorithm.

Definition 5.2

A one-point algorithm ϕ *is* optimal if $\phi \in \Phi_{n^*}$, has order n^* and has minimal combinatory cost. ■

Let d_3^* be the minimal combinatory cost of combining $\{F(x_i), F'(x_i), F''(x_i)\}$ to produce z_{i+1} such that $z_{i+1} - \alpha = O(e_i^3)$ in the scalar case. Let $d_2^*(N)$ be the minimal combinatory cost of combining $\{F(x_i), F'(x_i)\}$ to produce z_{i+1} such that $\|z_{i+1} - \alpha\| = O(e_i^2)$ in the multivariate case. We know

$$3 \le d_3^* \le 7$$

$$N+N^2 \le d_2^*(N) = O(N^\beta) \quad \text{where } \beta = \lg 7.$$

Recall that x_{i+1}^* is a zero of the interpolatory polynomial $w_i(x)$. Then $\|x_{i+1}^* - z_{i+1}\| = O(e_i^p)$ where $p = 3$ for $N = 1$ and $p = 2$ for $N \ge 2$. Thus d_3^* is the complexity of approximating a zero of a scalar quadratic equation and $d_2^*(N)$ is the complexity of solving a linear equation of order N.

Using the iteration I_3^* with an algorithm of minimal complexity for approximating a zero of a scalar quadratic equation and Newton iteration $IN_2^* = I_2^*$ with an algorithm of minimal complexity for solving of linear equations, we get from theorem 5.2

$$z(IN_3^*) = z_3(F) = z(F) \quad \text{for } N = 1$$
$$z(IN_2^*) = z_2(F) = z(F) \quad \text{for } N \ge 2.$$

Thus we have

Theorem 5.3

The iteration I_3^* is optimal for the scalar case and Newton iteration I_2^* is optimal for the multivariate case among all one-point iterations whenever F satisfies the equal cost assumption and $c(F) \ge H_N$ (as defined in theorem 5.2). ■

Remark 5.2

Along the complexity dimension Newton iteration is optimal in the multivariate case. But along the convergence dimension the iterations I_n with large n seem to be more attractive. However, to preserve a "type of global convergence" of I_n one has to approximate a zero of a polynomial operator equation and this decreases the radius of convergence (see Traub and Woźniakowski [76c]). ■

We have established the informational optimality of IN_3 and IN_2 with respect to the complexity index. Of course our primary interest is in complexity. Combining theorems 5.1 and 5.2 we obtain

Theorem 5.4

If $c(F^{(i)}) = \binom{N+i-1}{i} c(F)$ for $i \ge 1$ and

(i) for $N = 1$, $\underline{G}(3)(t\Gamma)^2 \ge 1$ and $\bar{G}(3)(2\Gamma)^2 \le 1$ then

$$\left(\frac{3}{\lg 3} c(F) + \frac{7}{\lg 3}\right)(\lg t - \lg\lg t) \le \mathrm{comp}(IN_3) \le \left(\frac{3}{\lg 3} c(F) + \frac{7}{\lg 3}\right) \lg(1+t)$$

(ii) for $N \ge 2$, $\underline{G}(2)(t\Gamma) \ge 1$ and $\bar{G}(2)(2\Gamma) \le 1$ then

$$(N+1)[c(F) + d(IN_2)/(N+1)](\lg t - \lg\lg t) \le \mathrm{comp}(IN_2) \le (N+1)[c(F) + d(IN_2)(N+1)] \cdot$$
$$\cdot \lg(1+t).$$ ■

6. OPTIMAL INFORMATION NUMBER FOR A SPECIAL PROBLEM

If F does not satisfy the equal cost assumption then the optimal information number may be large. For illustration we consider the class of problems of fixed dimension N where the cost c(F) of evaluating F varies and each piece of data in $F'(x), F''(x), \ldots$ costs the same number of arithmetic operations (for instance if F is an integral this often holds). This means that

$$(6.1)\quad c(F^{(i)}) = \binom{N+i-1}{i} c_1 \quad \text{for } i = 1,2,\ldots$$

where $c_1 = c_1(N)$ is a positive constant. Then the information complexity is given by

$$u(F;\mathfrak{N}_n) = \left\{\binom{N+n-1}{n-1} - 1\right\} c_1 + c(F).$$

From (4.19), (5.8) and (5.7) the minimal complexity index is bounded by

$$(6.2)\quad \frac{c(F) + (N+c_1)\binom{N+n-1}{n-1} - c_1}{\lg n} \le z_n(F) \le \frac{c(F) + \left\{\binom{N+n-1}{n-1} - 1\right\} c_1 + KN^2\binom{N+n-1}{n-1} \cdot \lg n}{\lg n}$$

for a positive constant K independent of N and n.

Recall that we define n^* by

$$(6.3)\quad z(F) = z_{n^*}(F) = \min_{n\ge 2} z_n(F).$$

The minimum is achieved since $z_n(F)$ tends to infinity with n. We are interested in finding n^* for large c(F). After some tedious algebraic manipulations we have

Theorem 6.1

$$(6.4)\quad z_{n^*}(F) = N\,\frac{c(F)}{\lg c(F)}(1+o(1)),\ \text{as } c(F) \to +\infty\ .$$

For every $\gamma > 1$ there exists $c_0 = c_0(\gamma)$ such that

$$(6.5)\quad c(F)^{\frac{1}{N\gamma}} \le n^* \le c(F)^{\frac{\gamma}{N}} \quad \text{for } c(F) \ge c_0.$$

■

Theorem 6.1 states that the minimal complexity index is roughly equal to $N\,c(F)/\lg c(F)$ and the optimal information number n^* tends to infinity almost linearly with $c(F)^{\frac{1}{N}}$. Note that

$$z_n(F) = \frac{c(F)}{\lg n}\,(1+o(1)) \quad \text{as } c(F) \to +\infty$$

for fixed n. This means that the "penalty" associated with using non-optimal fixed n is equal to

$$(6.6)\quad \frac{z_n(F)}{z_{n^*}(F)} = \frac{1}{N}\lg c(F)(1+o(1))$$

which tends to infinity with $c(F)$.

REFERENCES

Brent [76] — Brent, R. P., "A Class of Optimal Order Zero Finding Methods Using Derivative Evaluation," in Analytic Computational Complexity, edited by J. F. Traub, Academic Press, 1976.

Brent and Kung [76] — Brent, R. P. and Kung, H. T., "Fast Algorithms for Manipulating Formal Power Series," Dept. of Computer Science Report, Carnegie-Mellon University, 1976.

Kung and Traub [74] — Kung, H. T. and Traub, J. F., "Computational Complexity of One-Point and Multi-Point Iteration," Complexity of Computation, edited by R. Karp, AMS, 1974, 149-160.

Ortega and Rheinboldt [70] — Ortega, J. M. and Rheinboldt, W. C., Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

Rall [69] — Rall, L. B., Computational Solution of Nonlinear Operator Equations, John Wiley and Sons, New York, 1969.

Traub [64] — Traub, J. F., Iterative Methods for the Solution of Equations, Prentice-Hall, 1964.

Traub and Woźniakowski [76a] — Traub, J. F. and Woźniakowski, H., "Strict Lower and Upper Bounds on Iterative Computational Complexity," in Analytic Computational Complexity, edited by J. F. Traub, Academic Press, 1967, 15-34.

Traub and Woźniakowski [76b] — Traub, J. F. and Woźniakowski, H., "Optimal Linear Information for the Solution of Non-linear Equations," in Algorithms and Complexity: New Directions and Recent Results, edited by J. F. Traub, Academic Press, 1976.

Traub and Woźniakowski [76c] — Traub, J. F. and Woźniakowski, H., "Optimal Radius of Convergence of Interpolatory Iterations for Operator Equations," Dept. of Computer Science Report, Carnegie-Mellon University, 1976.

Traub and Woźniakowski [77] — Traub, J. F. and Woźniakowski, H., "Convergence and Complexity of Newton Iteration for Operator Equations," Dept. of Computer Science Report, Carnegie-Mellon University, 1977.

Woźniakowski [75a]	Woźniakowski, H., "Generalized Information and Maximal Order of Iteration for Operator Equations," SIAM J. Numer. Anal., Vol. 12, No. 1, March 1975, 121-135.
Woźniakowski [75b]	Woźniakowski, H., "Numerical Stability for Solving Nonlinear Equations," Dept. of Computer Science Report, Carnegie-Mellon University, 1975. To appear in Num. Math.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CONVERGENCE AND COMPLEXITY OF INTERPOLATORY-NEWTON ITERATION IN A BANACH SPACE		5. TYPE OF REPORT & PERIOD COVERED Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J. F. Traub and H. Wozniakowski		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0370 NR 044-422
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Dept. Pittsburgh, PA 15213		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE April 1977
		13. NUMBER OF PAGES 33 (12) 34 p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The class of Interpolatory-Newton iterations is defined and analyzed for the computation of a simple zero of a non-linear operator in a Banach space of finite or infinite dimension. Convergence of the class is established. The concepts of "informationally optimal class of algorithms" and "optimal algorithm" are formalized. For the multivariate case, the optimality of Newton iteration is established in the class of one-point iterations under an "equal cost assumption".

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

403 081